PAGE: 1

1652 #8

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/084,491A*

DATE: 05/28/1999
TIME: 11:35:52

INPUT SET: S32049.raw

> **This Raw Listing contains the General
> Information Section and up to the first 5 pages.**

```
 1                                SEQUENCE LISTING        ENTERED
 2
 3   (1)     General Information:
 4
 5        (i) APPLICANT: MOORE, PAUL A.
 6   RUBEN, STEVEN M.
 7   EBNER, REINHARD
 8
 9  ·     (ii) TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
10
11      (iii) NUMBER OF SEQUENCES: 16
12
13       (iv) CORRESPONDENCE ADDRESS:
14            (A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
15            (B) STREET: 9410 KEY WEST AVENUE
16            (C) CITY: ROCKVILLE
17            (D) STATE: MD
18            (E) COUNTRY: USA
19            (F) ZIP: 20850
20
21        (v) COMPUTER READABLE FORM:
22            (A) MEDIUM TYPE: Floppy disk
23            (B) COMPUTER: IBM PC compatible
24            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
25            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
26
27       (vi) CURRENT APPLICATION DATA:
28            (A) APPLICATION NUMBER: 09/084,491
29            (B) FILING DATE: 27-MAY-1998
30            (C) CLASSIFICATION:
31
32     (viii) ATTORNEY/AGENT INFORMATION:
33            (A) NAME: BROOKES, ANDERS A.
34            (B) REGISTRATION NUMBER: 36,373
35            (C) REFERENCE/DOCKET NUMBER: PF378
36
37       (ix) TELECOMMUNICATION INFORMATION:
38            (A) TELEPHONE: (301) 309-8504
39            (B) TELEFAX: (301) 309-8439
40
41
42   (2) INFORMATION FOR SEQ ID NO:1:
43
44        (i) SEQUENCE CHARACTERISTICS:
45            (A) LENGTH: 2329 base pairs
46            (B) TYPE: nucleic acid
```

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/084,491A*

DATE: 05/28/1999
TIME: 11:35:52

INPUT SET: S32049.raw

```
47              (C) STRANDEDNESS: single
48              (D) TOPOLOGY: linear
49
50      (ii) MOLECULE TYPE: DNA (genomic)
51
52
53      (ix) FEATURE:
54              (A) NAME/KEY: CDS
55              (B) LOCATION: 124..913
56
57      (ix) FEATURE:
58              (A) NAME/KEY: sig_peptide
59              (B) LOCATION: 124..184
60
61      (ix) FEATURE:
62              (A) NAME/KEY: mat_peptide
63              (B) LOCATION: 187..913
64
65
66
67      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
68
69   TTACCAGAAC AGCATAACAA GGGCAGGTCT GACTGCAAGC TGGGACTGGG AGGCAGAGCC        60
70
71   GCCGCCAAGG GGGCCTCGGT TAAACACTGG TCGTTCAATC ACCTGCAAGA CGAAGAGGCA       120
72
73   AGG ATG CTG TTG GCC TGG GTA CAA GCA TTC CTC GTC AGC AAC ATG CTC       168
74       Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu
75       -21 -20              -15              -10
76
77   CTA GCA GAA GCC TAT GGA TCT GGA GGC TGT TTC TGG GAC AAC GGC CAC       216
78   Leu Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His
79        -5                  1                5                   10
80
81   CTG TAC CGG GAG GAC CAG ACC TCC CCC GCG CCG GGC CTC CGC TGC CTC       264
82   Leu Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu
83                15                   20                   25
84
85   AAC TGG CTG GAC GCG CAG AGC GGG CTG GCC TCG GCC CCC GTG TCG GGG       312
86   Asn Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly
87                30                   35                   40
88
89   GCC GGC AAT CAC AGT TAC TGC CGA AAC CCG GAC GAG GAC CCG CGC GGG       360
90   Ala Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly
91                45                   50                   55
92
93   CCC TGG TGC TAC GTC AGT GGC GAG GCC GGC GTC CCT GAG AAA CGG CCT       408
94   Pro Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro
95        60                   65                   70
96
97   TGC GAG GAC CTG CGC TGT CCA GAG ACC ACC TCC CAG GCC CTG CCA GCC       456
98   Cys Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala
99   75                   80                   85                   90
```

INPUT SET: S32049.raw

```
100
101   TTC ACG ACA GAA ATC CAG GAA GCG TCT GAA GGG CCA GGT GCA GAT GAG      504
102   Phe Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu
103                   95                  100                 105
104
105   GTG CAG GTG TTC GCT CCT GCC AAC GCC CTG CCC GCT CGG AGT GAG GCG      552
106   Val Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala
107               110                 115                 120
108
109   GCA GCT GTG CAG CCA GTG ATT GGG ATC AGC CAG CGG GTG CGG ATG AAC      600
110   Ala Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn
111               125                 130                 135
112
113   TCC AAG GAG AAA AAG GAC CTG GGA ACT CTG GGC TAC GTG CTG GGC ATT      648
114   Ser Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile
115       140                 145                 150
116
117   ACC ATG ATG GTG ATC ATC ATT GCC ATC GGA GCT GGC ATC ATC TTG GGC      696
118   Thr Met Met Val Ile Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly
119   155                 160                 165                 170
120
121   TAC TCC TAC AAG AGG GGG AAG GAT TTG AAA GAA CAG CAT GAT CAG AAA      744
122   Tyr Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys
123                   175                 180                 185
124
125   GTA TGT GAG AGG GAG ATG CAG CGA ATC ACT CTG CCC TTG TCT GCC TTC      792
126   Val Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe
127                   190                 195                 200
128
129   ACC AAC CCC ACC TGT GAG ATT GTG GAT GAG AAG ACT GTC GTG GTC CAC      840
130   Thr Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His
131           205                 210                 215
132
133
134   ACC AGC CAG ACT CCA GTT GAC CCT CAG GAG GGC AGC ACC CCC CTT ATG      888
135   Thr Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met
136       220                 225                 230
137
138   GGC CAG GCC GGG ACT CCT GGG GCC T GAGCCCCCCC AGTGGGCAGG            933
139   Gly Gln Ala Gly Thr Pro Gly Ala
140   235                 240
141
142   AGCCCATGCA GACACTGGTG CAGGACAGCC CACCCTCCTA CAGCTAGGAG GAACTACCAC    993
143
144   TTTGTGTTCT GGTTAAAACC CTACCACTCC CCCGCTTTTT TGGCGAATCC TAGTAAGAGT    1053
145
146   GACAGAAGCA GGTGGCCCTG TGGGCTGAGG GTAAGGCTGG GTAGGGTCCT AACAGTGCTC    1113
147
148   CTTGTCCATC CCTTGGAGCA GATTTTGTCT GTGGATGGAG ACAGTGGCAG CTCCCACAGT    1173
149
150   GATGCTGCTG CTAAGGGCTT CCAAACATTG CCTGCACCCC TGGAACTGAA CCAGGGATAG    1233
151
152   ACGGGGAGCT CCCCCAGGCT CCTCTGTGCT TTACTAAGAT GGCTCAGTCT CCACTGTGGG    1293
```

INPUT SET: S32049.raw

```
153
154   CTTGAGTGGC ATACACTGTT ATTCATGGTT AAGGTAAAGC AGGTCAAGGG ATGGCATTGA    1353
155
156   AAAAATATAT TTAGTTTTTA AAATATTTGG GATGGAACTC CCTACTGACC TCTGACAACT    1413
157
158   GGAAACGAGT TTGTACTGAA GTCAGAACTT TGGGTTGGGA ATGAGATCTA GGTTGTGGCT    1473
159
160   GCTGGTATGC TTCAGCTTGC TGGCAATGAT GTGCCTTGAC AACCGTGGGC CAGGCCTGGG    1533
161
162   CCCAGGGACT CTTCCTGTTT CATAAGGAAA GGAAGAATTG CACTGAGCAT TCCACTTAGG    1593
163
164   AAGAGGATAG AGAAGGATCT GCTCCGCCTT TGGCCACAGG AGCAGAGGCA GACCTGGGAT    1653
165
166   GCCCCAGTTT CTCTTCAGGG ATGGATAGTG ACCTGTCTTC ATTTTGCACA GGTAAGAGAG    1713
167
168   TAGTTAGCTA ACCTATGGGA ATTATACTGT GGGGCCTTGT GAGCTGCTTC TAAGAGGCTA    1773
169
170   ACCTGGAAAC TAAGCTCAGA GGCAAGGTAA TAAAGCACTT CAGGGCTTGC TCCCCAAGTG    1833
171
172   GGCCTGATTT AGCAGGTGGT CTGCGGGCGT CCAGGTCAGC ACCTTCCTGT AGGGCACTGG    1893
173
174   GGCTAGGGTC ACAGCCCCTA ACTCATAAAG CAATCAAAGA ACCATTAGAA AGGGCTCATT    1953
175
176   AAGCCTTTTG GACACAGGAC CCCAGAGAGG AAAAAGTGAC TTGCCCAAGG TCGTAAGCAA    2013
177
178   GCTACTGGCA TGGCAAGAGC CCAGCTTCCT GACGGAGCGC AACATTTCTC CACTGCACTG    2073
179
180   TGCTAGCAGC TCAGCAGGGC CTCTAACCTG TGATGTCACA CTCAAGAGGC CTTGGCAGCT    2133
181
182   CCTAGCCATA GAGCTTCCTT TCCAGAACCC TTCCACTGCC CAATGTGGAG ACAGGGGTTA    2193
183
184   GTGGGGCTTT CTATGGAGCC ATCTGCTTTG GGGACCTAGA CCTCAGGTGG TCTCTTGGTG    2253
185
186   TTAGTGATGC TGGAGAAGAG AATATTACTG GTTTCTACTT TTCTATAAAG GCATTTCTCT    2313
187
188   ATAAAAAAAA AAAAAA                                                    2329
189
190
191   (2) INFORMATION FOR SEQ ID NO:2:
192
193           (i) SEQUENCE CHARACTERISTICS:
194               (A) LENGTH: 263 amino acids
195               (B) TYPE: amino acid
196               (D) TOPOLOGY: linear
197
198           (ii) MOLECULE TYPE: protein
199
200
201           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
202
203   Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu Leu
204   -21 -20              -15                 -10
205
```

INPUT SET: S32049.raw

```
206  . Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His Leu
207    -5                  1               5                10
208

209    Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu Asn
210            15              20                  25
211

212    Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly Ala
213        30              35                  40
214

215    Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly Pro
216        45              50                  55
217

218    Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro Cys
219    60              65                  70                  75
220

221    Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala Phe
222            80                  85                  90
223

224  - Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu Val
225            95                  100                 105
226

227    Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala Ala
228            110                 115                 120
229

230    Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn Ser
231        125                 130                 135
232

233    Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile Thr
234    140             145                 150                 155
235

236    Met Met Val Ile Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly Tyr
237            160                 165                 170
238

239    Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys Val
240            175                 180                 185
241

242    Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe Thr
243            190                 195                 200
244

245    Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His Thr
246        205                 210                 215
247

248    Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met Gly
249    220                 225                 230                 235
250

251    Gln Ala Gly Thr Pro Gly Ala
252                240
253

254    (2) INFORMATION FOR SEQ ID NO:3:
255

256        (i) SEQUENCE CHARACTERISTICS:
257            (A) LENGTH: 372 amino acids
258            (B) TYPE: amino acid
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/09/084,491A*

INPUT SET: S32049.raw

Line	Error	Original Text